Annals of R.S.C.B., ISSN:1583-6258, Vol. 25, Issue 4, 2021, Pages. 9656 - 9666
Received 05 March 2021; Accepted 01 April 2021.

Content Based Image Retrieval System Using Integrated ML and DL-CNN

Mounika Jammula
Assistant Professor, Department of Electronics and Communication Engineering
Chaitanya Bharathi Institute of Technology, Hyderabad, India
E-mail: jmounika_ece@cbit.ac.in

Abstract

In this article, content-basedimage retrieval (CBIR) system is developed since it has significant scope for research in image processing domain. In CBIR, visual contents are being used to search an image from a huge scale image database as per users interests and based on automatically derived query image features. The term 'content' might refer to low level features like color, shape or texture extracted from the image. Several research challenges can be addressed towards the design and development of CBIR systems, very few techniques are available to address and solve the problem of semantic gap presented in images which are not efficient.Themachine learning (ML) method has investigated as a practicableapproach to decrease the semantic gap. Also motivatedfrommodernfulfillment of deep learning for image processing applications, we focused to deal with an artificial intelligence based deep learning approach, treated as Convolutional-Neural-Network (CNN), for the purpose ofsimilarity measurement of accuratesemantic features. In this article, the usage of CNN for the image retrieval issues is investigatedwith their solutions using deep learning approaches.Further, it is also incorporated with principal component analysis (PCA) for extracting salient features from the images. Euclidean distance measurement is employed for the similarity evaluation of extracted feature vectors of query image and database images. Extensive simulation results on different image categories discloses that proposed DL-CNN-ML outperforms existing CBIR approaches like ML and CNN in terms of mean average precision (AP), mean average recall (mAR) and F-score values.

Keywords: content-based image retrieval, feature extraction, machine learning, deep learning, principal component analysis and convolutional neural networks.

I. INTRODUCTION

In current days, there is a quick improvement in picture search situations, for example, Google's picture - search, Microsoft's CBIR innovation, Bing CBIR instrument, and note: does not deal with all pictures (Public Company), CBIR internet searcher, by Gazopa (Private Company), Imense Image Search Portal (Private Company) and Like.com (Private Company), picture recovery has become a difficult assignment. The enthusiasm for CBIR has developed in view of the recovery issues, confinements, and time utilization in metadata-based frameworks. We can look through the printed data effectively by the current innovation, however this looking through strategies expects people to portray every picture physically in the database, which is unimaginable for all intents and purposes for exceptionally gigantic databases or for the pictures which will be produced consequently, for example pictures produced from reconnaissance cameras. It has more disadvantages that there is an opportunity to miss pictures that utilization distinctive proportionate word in the depiction of pictures. The frameworks dependent on arranging pictures in semantic classes like "tiger" as a subclass of "creature" can suspend the miscatergorization issue, however it will require more exertion by an utilization to recognize the pictures that may be "tigers", yet every one of them are sorted distinctly as a "creature". CBIR is a use of techniques for obtaining, pre-preparing, investigating, portrayal and furthermore understanding pictures to the picture recovery issue, that is the issue of investigating for advanced pictures from enormous databases. The CBIR framework is against conventional methodologies, which is

Annals of R.S.C.B., ISSN:1583-6258, Vol. 25, Issue 4, 2021, Pages. 9656 - 9666
Received 05 March 2021; Accepted 01 April 2021.

known on context-based methodologies i.e., context-based image (CBII) [1].Representation of highlights and likeness estimations are basic for the recovery execution of a CBIR framework. Different methodologies have been proposed, however and still, after all that, it stays as a difficult assignment because of the semantic hole present between the picture pixels and significant level semantics saw by people. One ideal methodology is ML that means to take care of this issue in the long haul. Profound learning speaks to a class of ML approaches where a few layers of information preparing steps in various leveled designs are used for characterization errand and investigation of highlights [2]. Profound learning systems have accomplished incredible accomplishments in picture arrangement. In any case, the positioning of comparative pictures is conflicting with the grouping of pictures. For arrangement of pictures, "dark boots," "white boots" and "dull dim boots" are generally boots, yet for positioning of comparable pictures, if a question picture is a "dark boot," we customarily need to rank the "dim boot" higher than the "white boot." CNNs [2] are a particular sort of ANN for taking care of information that includes a framework like topology like, picture information, which is a 2D network of pixels. CNNs are just ANNs that includes the utilization of convolution rather than traditional framework augmentation activity in at least one in the entirety of their layers. Convolution bolsters three fundamental ideas that can encourage in improving a ML framework: parameter sharing, equivariant portrayals, and scanty collaborations. CNNs are prominent for their capability to learn shapes, surfaces, and hues, making this issue reasonable for the utilization of neural systems.

In this, we investigated an architecture of integrated ML with DL-CNN for CBIR systems. Initially, the training operation will be performed by using DL-CNN and PCA on the database to create the feature vectors, and the fully connected layer holds all the feature vectors. When the test image is applied, then its feature vector is compared, and similarity will be retained by using Euclidean distance among those feature vectors.

II.RELATED WORK

There are many researchers, who have published numerous articles on CBIR system using deep learning models. Authors in [5], trained a deep CNN framework for classifying a dataset of ImageNet comprising of the images approximately 1.2millions into number of classes (around 1000), where the authors utilized 8-layered network with first five of convolutional and later are fully connected layers. They have utilized the features extracted form the seventh layer to obtain similar images and accomplished 37.5% error rate fortop-1 and 17% for top-5. But CNN features have higher dimensionality and unskillfulness of resemblance calculation between couple of vectors with 4096 dimensions. Later, mitigation of dimensionality has been presented to reduce the features dimensionality [6].Study of binary has codes using an approach called supervised hashing was proposed in [7], which utilized deep learning models for image retrieval system and disclosed superior performance evaluation on publicly available datasets. However, the preprocessing stage is exceptionally critical if there is a large data to be processed as it consumes high storage and higher computational time. Similarly, supervised learning-based hashing technique has been proposed by Lin et al. [8] for faster image retrieval system, where the authors converted the features space of high-dimensional into low-dimensional and rendered the hash codes of binary format. In addition, a binary pattern matching concept also employed which amazingly mitigated the computational time and optimizes the efficaciousness of image searching from the database. Further, they have utilized hamming distance metric instead of traditional Euclidean distance for computational complexity reduction.Authorsin [3] has introduced a firefly neural network to classify the CBIR of hand, chest and back. Based on fuzzy SVM classifier using a firefly neural network to classify the CBIR of hand, chest and back. The

Annals of R.S.C.B., ISSN:1583-6258, Vol. 25, Issue 4, 2021, Pages. 9656 - 9666
Received 05 March 2021; Accepted 01 April 2021.

major advantage of the proposed system is performance and accuracy improvements include form-based function, Active Appearance Model and Trace surface texture. The major limitation of the proposed system is back images are not classifiedcorrectly; some noises are interrupted.Wang et al. [9] have studied image segmentation, extraction and representation of feature, index, and related retrieval techniques. Aiming at image segmentation, a KNN based CBIR method based on canny edge detection is proposed. Based on the traditional gray histogram feature extraction technique, an adaptive weighted improved gray histogram method is proposed, and it is proved by experiments that this method can enhance some important features of the image. It is easier to calculate the similarity and help the doctor to find the image features of interest in the complex learning image. Vikram M et al. [10] has introduced a LDA based method for images encoding the visual features. To evaluate the proposed system the author has considered with the help of Image CLEF 2009 dataset. The major advantage of the introduced method is Method used to efficiently signify the topics of the clustered features of SIFT. The proposed method archives 0.32 of Mean Average Precision and 0.722 of Normalized Discounted Cumulative Gain. The major limitation of the introduced system is Geometric Mean Average Precision value is low.Dash JK et al. [11] has proposed a Multiple Classifier System used for texture image retrieval similarity. To evaluate the proposed system the authors have considered USC-SIPI Image Database. And the author has split into the dataset as a D1, D2, D3 and D4. The proposed system archive Retrieval performance 96.72 % on database D1, and 95.26 % on database D2, and 91.11 % on database D3 and 83.64 % on database D4, the major advantage of the introduced method is Higher SLSC value signifies lower execution period. But Performance of Computation of feature is not better.

III. PROPOSED IMPLEMENTATION

This subsection discussed about the proposedapproach which utilizesDL-CNN-ML for CBIR implementation as disclosed in Figure 1. Methodology of CNN implemented uses 2-Dimensional convolutional layer creates and apply the movable filters to original test input image.This layer convolutes the source imageby changing filter size and changing their position with alongthe horizontal and vertical weights then perform dot product operationbetween source images to their corresponding weights,finally adds the bias term for decision taking in the layers. A rectified linear unit layer of CNN executes a thresholdingprocess to everypixel element and their property of test image, if any pixel value is lessthan threshold value set to zero(0). Then max-pooling-layer divides the input data into multiple rectangular pooling parts by using down-samplingand calculates the highestdensity of everypart. Then fully-connected-layer performs do multiplication operation between the sourceimagesto its weight matrix feature and finally adds the bias term for decision taking vector.

Annals of R.S.C.B., ISSN:1583-6258, Vol. 25, Issue 4, 2021, Pages. 9656 - 9666
Received 05 March 2021; Accepted 01 April 2021.

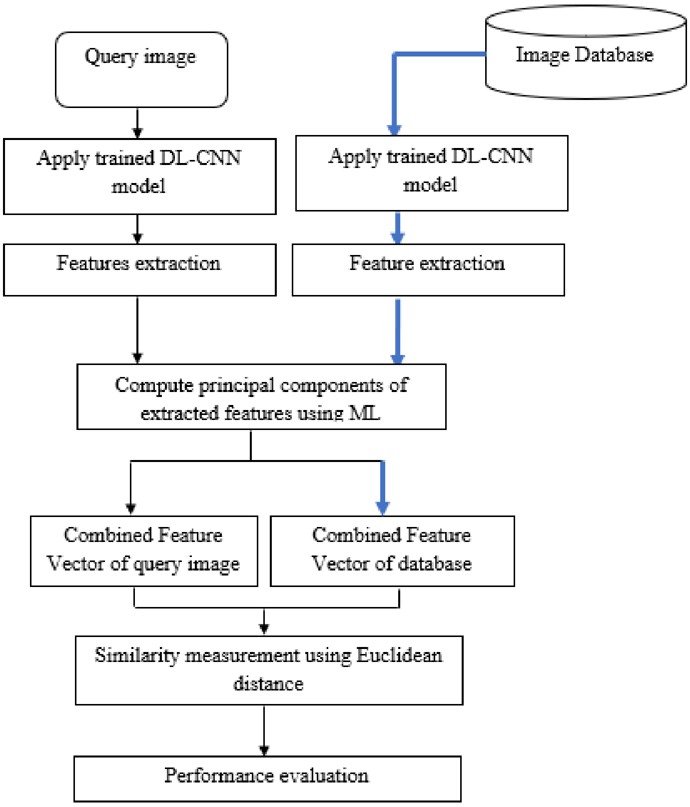


Fig. 1: Proposed CBIR approach using DL-CNN-ML.

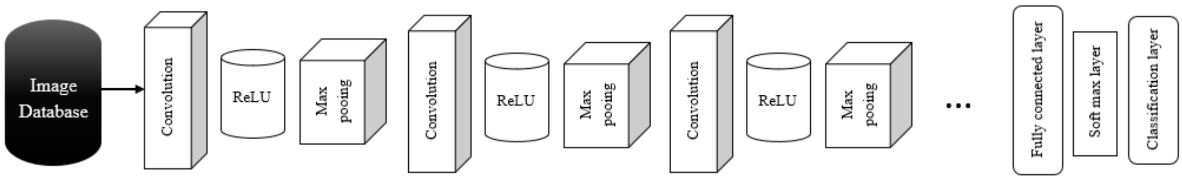


Fig. 2: Proposed DL-CNN model.

3.1. DL-CNN

As perthe essentials, testing and training of DL-CNN occupiesintoallocatingeachinputimage by the sequence of convolutional layers such as RELU layer, fully-connected-layer, max-pooling layer and SoftMaxlayer. Each layer consistingof most essential filter and kernel to implement the classification operation via parallel operation of layers to identify the objects and items using the probability density function which holds the probabilistic assessmentssortedfrom[0,1]. Thedesign of DL-CNN shown in Figure 2, that is used to identify the similar imagesfor the proposed CBIR approach withimprovedcharacteristicfeature indication and matching over novel image retrieval schemes.

3.2. Principal component analysis (PCA)

PCA is anefficient MLmethod that is used to decrease the area specificity to its dimensions. It usesfundamental matrices operationscommencing slinear morphological algebraand characteristics to calculateanoutcomefromprimaryinformation; the outcome contains lower dimensions but the comparable pixel count to that of input. PCA can be considered as a resultant method,where informationcontainsm-features or columns are expectedto be keen onanassociate-space through*m*number or slighterfeatures,while maintaining the important as well as majorityelement of original information. Let *I*ismatrix of anoriginal input image with a dimension of $n * m$ and outcomesoutput image*J*. The principal

Annals of R.S.C.B., ISSN:1583-6258, Vol. 25, Issue 4, 2021, Pages. 9656 - 9666
Received 05 March 2021; Accepted 01 April 2021.

step is to calculate the rate of mean for each columnfeature. After that, the values in eachfeature are created by deducting the original value from itsaveragefeature. Now, centered matrix is calculated with covariance properties.Finally, perform the decomposition operation on each covariance matrix property for eigenvaluecalculation thatprovides the group of eigenvectors. These eigenvectors comprise the information or apparatus or directionsfor the condensedassociate-space of J, while the maximumpotency for apparatus are indicated by eigenvectors. Normally, k eigenvectors would be chosen,they are referred as principal features or components.

3.3. Euclidean distance

To estimate distances amongtest image matrix I_q and resultant reclaimedimages I_r, a parameterhave to be calculated. We require a dimensiontechnique to inform how much test and resultant images are alike (pixel per pixel). Consequently, we desire a comparisonmetricin which the distance assessment will be number of comparablepixels in train images.

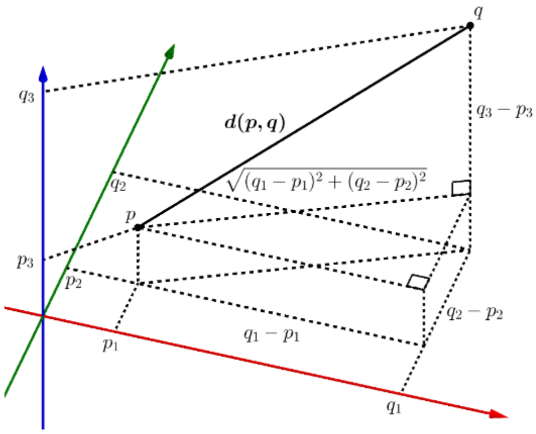


Fig. 3Representation of Euclidean distance

IV. RESULTS AND DISCUSSION

This section describes the implementation of proposed CBIR simulation results with comparison to the existing CBIR approaches. Figure 4 discloses the dataset images utilized as query images. Retrieved images of “Tiger” is shown in Figure 5, where existing ML-based CBIR approach and CNN-based CBIR approach is demonstrated in Figure 5(a) and Figure 5(b) respectively. Outcome of proposed DL-CNN-ML-based CBIR system is shown in Figure 5(c). From Figure 5, existing ML-based CBIR system failed to retrieve relevant images as given query image of “Tiger”. However, CNN-based CBIR system provides equal outcome with proposed DL-CNN-ML-based CBIR system for “Tiger” image as query, it is also failed to produce relevant images with “Dinosaur”, “Apple” and many other query images shown in Figure 4. Retrieved images of other images like “Dinosaur” and “Apple” is shown in Figure 6 and Figure 7 using existing and proposed CBIR systems.

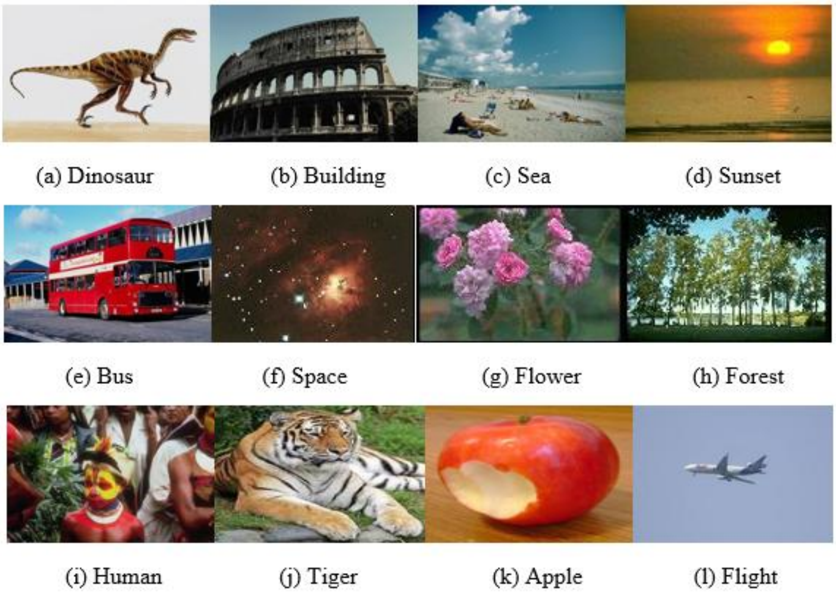
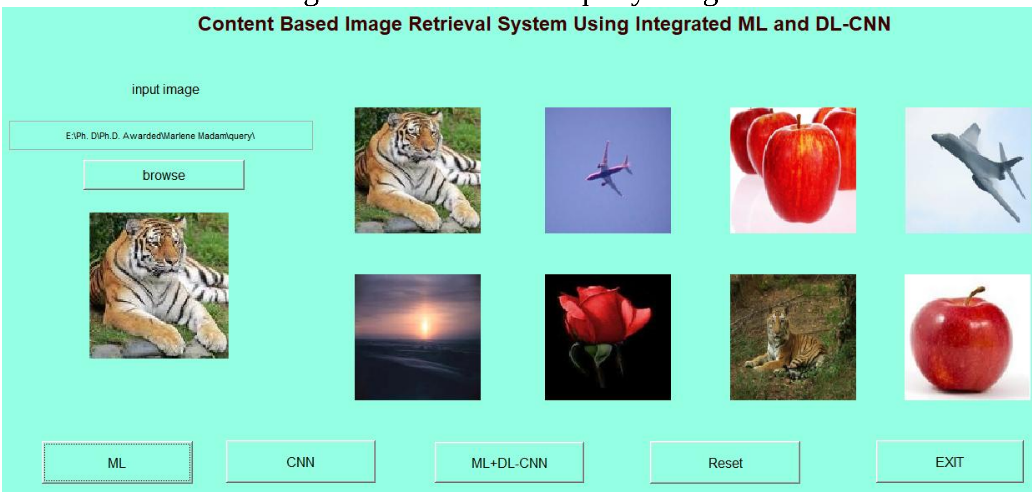
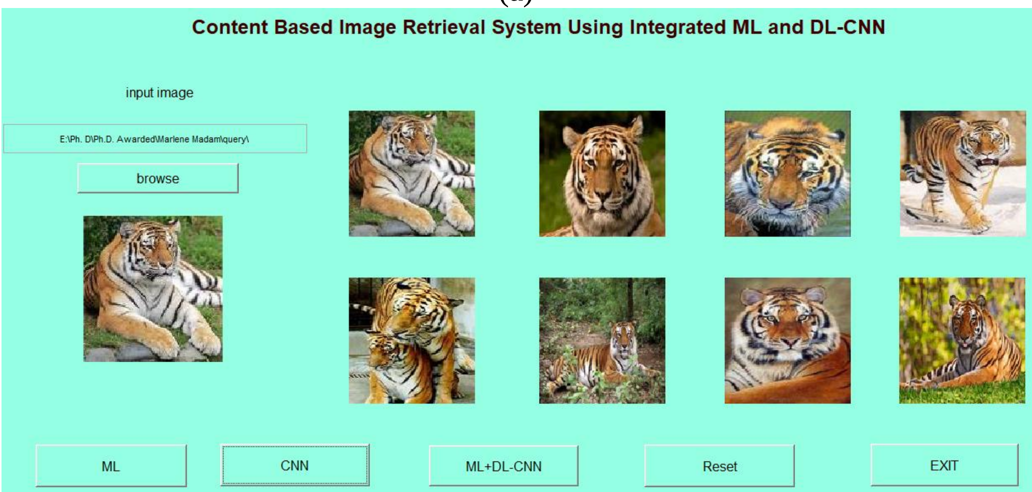


Fig. 4: Dataset used as query images.

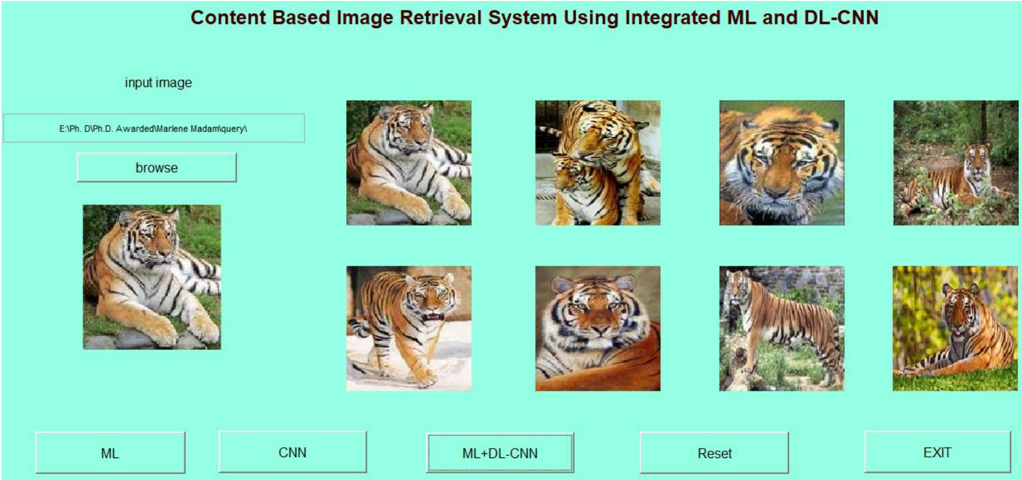


(a)



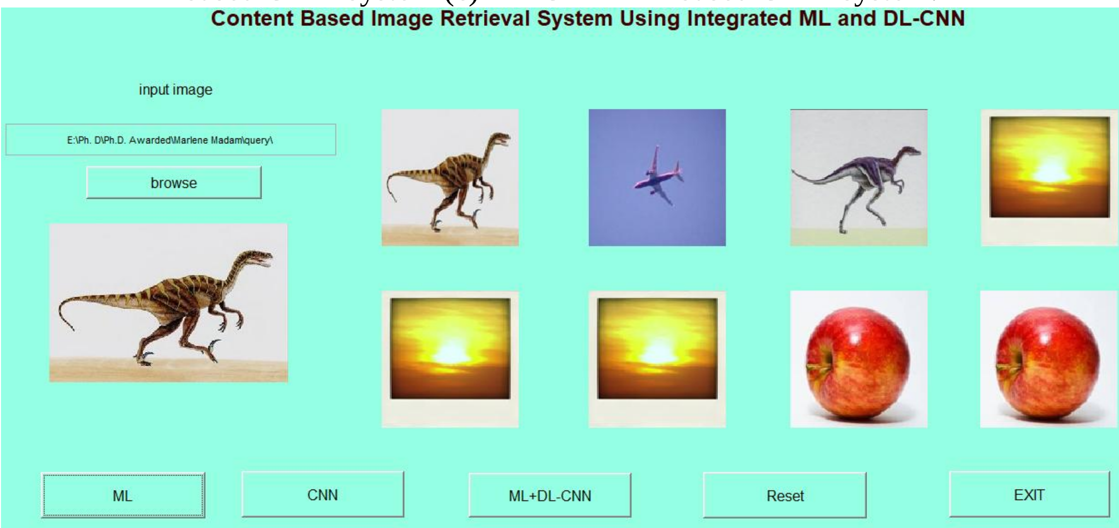
(b)

Annals of R.S.C.B., ISSN:1583-6258, Vol. 25, Issue 4, 2021, Pages. 9656 - 9666
Received 05 March 2021; Accepted 01 April 2021.

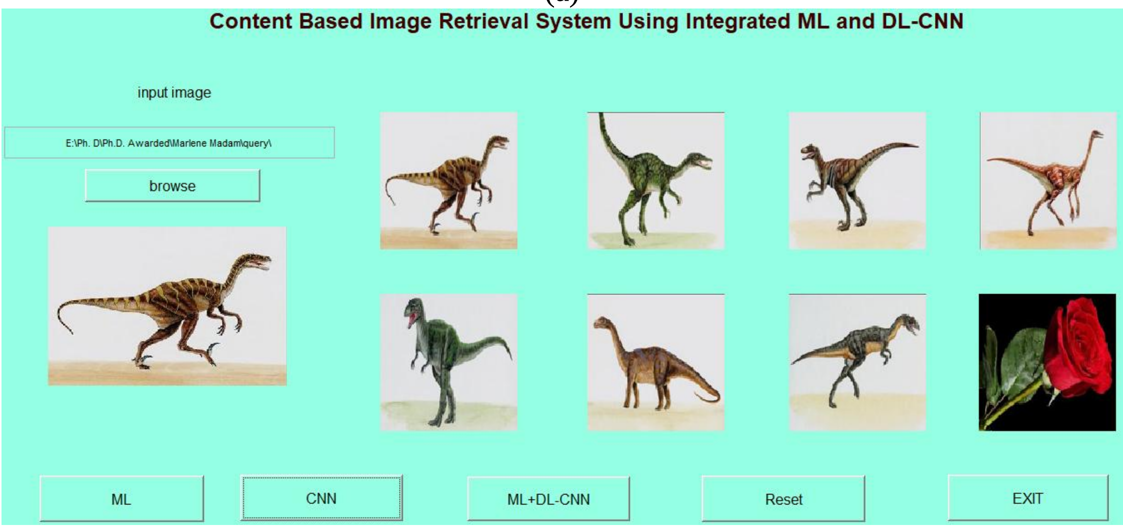


(c)

Fig. 5: Retrieved images of “Tiger” query image using (a) ML-based CBIR system (b) CNN-based CBIR system (c) DL-CNN-ML-based CBIR system.

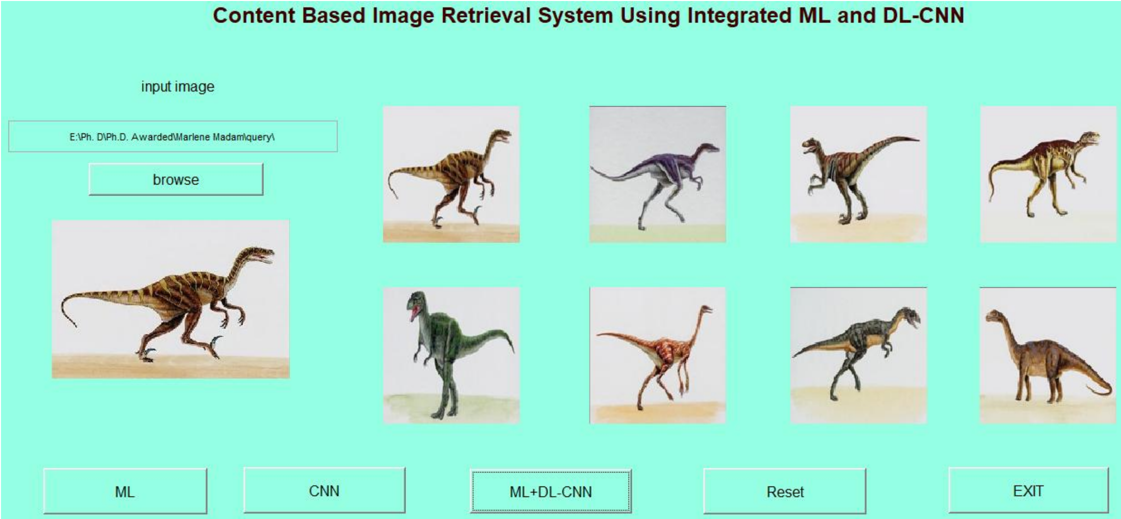


(a)



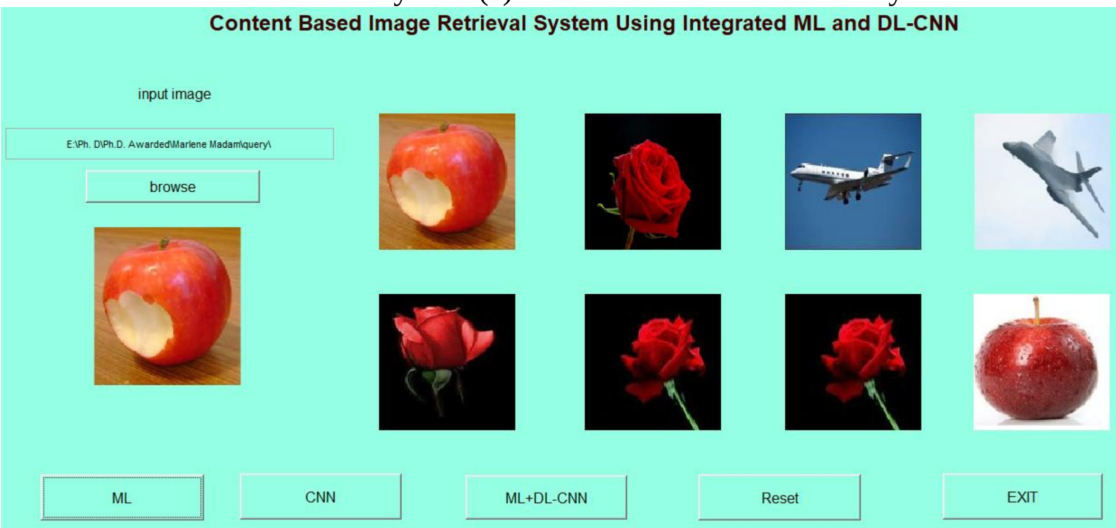
(b)

Annals of R.S.C.B., ISSN:1583-6258, Vol. 25, Issue 4, 2021, Pages. 9656 - 9666
Received 05 March 2021; Accepted 01 April 2021.

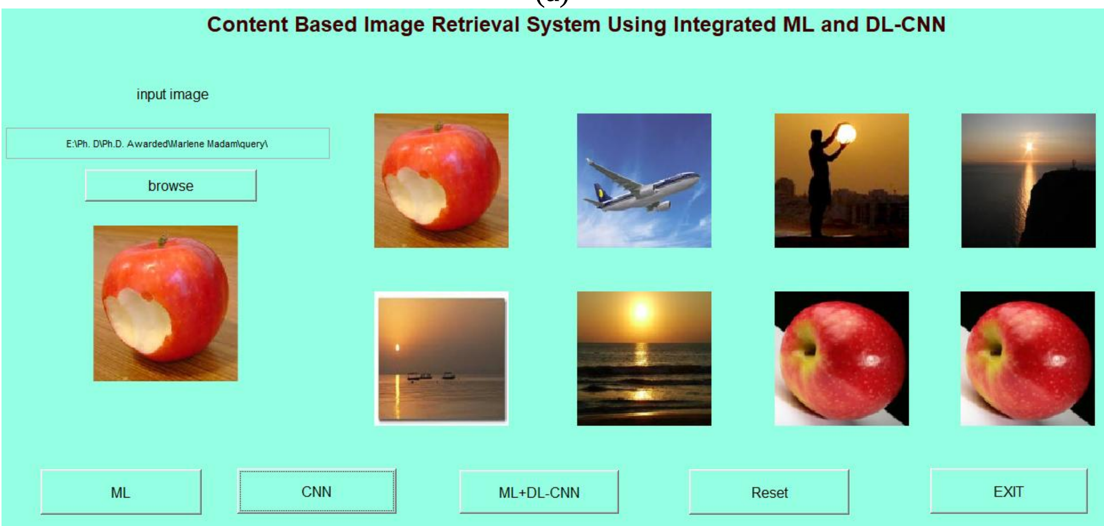


(c)

Fig. 6: Retrieved images of “Dinosaur” query image using (a) ML-based CBIR system (b) CNN-based CBIR system (c) DL-CNN-ML-based CBIR system.

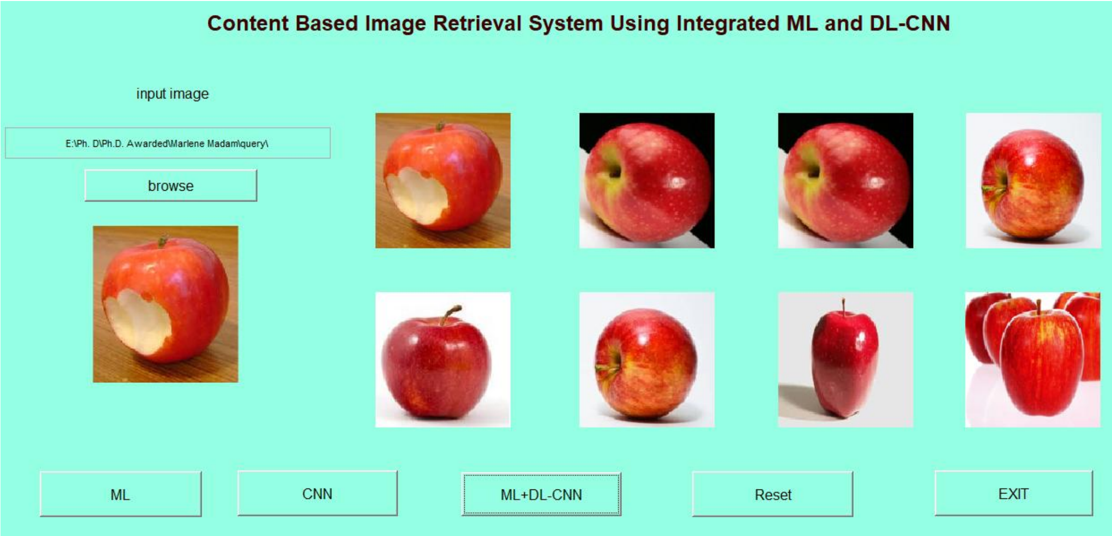


(a)



(b)

Annals of R.S.C.B., ISSN:1583-6258, Vol. 25, Issue 4, 2021, Pages. 9656 - 9666
Received 05 March 2021; Accepted 01 April 2021.



(c)

Fig. 7: Retrieved images of "Apple" query image using (a) ML-based CBIR system (b) CNN-based CBIR system (c) DL-CNN-ML-based CBIR system.

Quality evaluation

Precision is a metric for achieving capacity of CBIR method to recover merely significant imagery, and Recall defines the capability of CBIR method to recover the entire applicable imagery as definite by following equations correspondingly.

$$P = \frac{Total\ numebr\ \ of\ relevant\ \ images\ \ retrieved}{Total\ numebr\ \ of\ retrieved\ \ images} \qquad (1)$$

$$R = \frac{numebr\ \ of\ relevant\ \ images\ retrieved}{numebr\ \ of\ relevant\ \ images\ \ in\ database} \qquad (2)$$

F1-score is often used in the field of retrieval of information for measuring search, document classification, and query classification performance. Earlier works focused primarily on the F1-score, but with the proliferation of large-scale search engines, performance goals changed to place more emphasis on either precision or recall.

$$F = 2 \cdot \frac{\text{precision} \cdot \text{recall}}{\text{precision} + \text{recall}} \qquad (3)$$

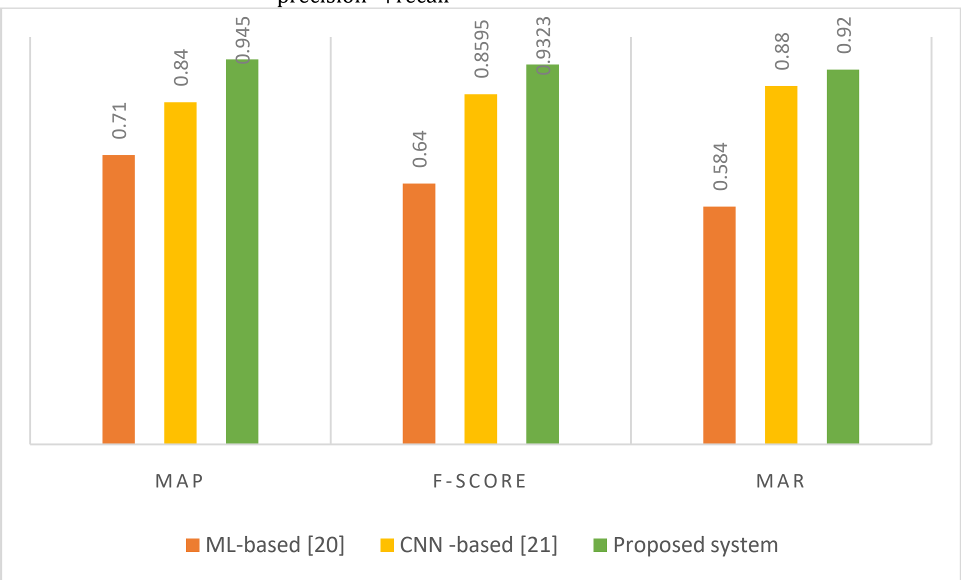


Fig. 8:Performance of mAP, F1-score, and mAR with proposed and existing CBIR systems.

Annals of R.S.C.B., ISSN:1583-6258, Vol. 25, Issue 4, 2021, Pages. 9656 - 9666
Received 05 March 2021; Accepted 01 April 2021.

V. CONCLUSION

This paperexplainsaproficient CBIR methodutilizingMLand DL-CNNwith pair wiseEuclidian distance matching. The implementationoutcomes shownthat proposed DL-CNN-ML-based CBIR system achievedbetter performance of accuracy by recallingfurthersimilarimagery. Additionally, the qualitativeperformance assessmentof proposed DL-CNN-ML-based CBIR systemis confirmedthat it outperforms existing ML-based and CNN-based CBIR systems in terms of mAR, mAP, and F-score.

REFERENCES

[1] Y. Liu, D. Zhang, G. Lu, and W.-Y. Ma, "A survey of content-based image retrieval with high-level semantics," Pattern recognition, vol. 40, no. 1, pp. 262–282, 2007.

[2] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," Nature, vol. 521, no. 7553, pp. 436–444, 2015.

[3] Rao TYS, Reddy PC. Content and context-based image retrieval classification based on firefly-neural network. Multimedia Tools and Applications. vol. 77, pp. 32041–32062, 2018.

[4] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, "Rethinking the inception architecture for computer vision," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016, pp. 2818– 2826.

[5] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in Advances in neural information processing systems, 2012, pp. 1097–1105.

[6] A. Babenko, A. Slesarev, A. Chigorin, and V. Lempitsky, "Neural codes for image retrieval," in European conference on computer vision. Springer, 2014, pp. 584–599.

[7] R. Xia, Y. Pan, H. Lai, C. Liu, and S. Yan, "Supervised hashing for image retrieval via image representation learning", in Proceedings of the AAAI Conference on Artificial Intelligence, vol. 1, 2014, pp. 2156-2162.

[8] K. Lin, H.-F. Yang, J.-H. Hsiao, and C.-S. Chen, "Deep learning of binary hash codes for fast image retrieval," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops, 2015, pp. 27–35.

[9] C. Wang, L. Liu, Y. Tan, "An Efficient Content-Based Image Retrieval System Using k-NN and Fuzzy Mathematical Algorithm," CMES-Computer Modeling in Engineering & Sciences, vol. 124, no. 3, pp. 1061–1083, 2020.

[10] M. Vikram, A. Anantharaman, B. S. Suhas,"An approach for multimodal medical image retrieval using latent Dirichlet allocation," In: Proceedings of the ACM India Joint International Conference on Data Science and Management of Data, pp. 44–51, 2019.

[11] Dash, J.K., Mukhopadhyay, S. Similarity learning for texture image retrieval using multiple classifier system. Multimedia Tools and Applications, vol. 77, pp. 459–483, 2018.

[12] N. Khosla and V. Venkataraman, "Building image-based shoe search using convolutional neural networks," CS231n Course Project Reports, 2015.

[13] A. Iliukovich-Strakovskaia, A. Dral, and E. Dral, "Using pre-trained models for fine-grained image classification in fashion field," 2016.

[14] J. Donahue, Y. Jia, O. Vinyals, J. Hoffman, N. Zhang, E. Tzeng, and T. Darrell, "Decaf: A deep convolutional activation feature for generic visual recognition." in Icml, vol. 32, 2014, pp. 647–655.

[15] G. Shrivakshan, C. Chandrasekar et al., "A comparison of various edge detection techniques used in image processing," IJCSI International Journal of Computer Science Issues, vol. 9, no. 5, pp. 272–276, 2012.

Annals of R.S.C.B., ISSN:1583-6258, Vol. 25, Issue 4, 2021, Pages. 9656 - 9666
Received 05 March 2021; Accepted 01 April 2021.

[16] J. Shlens, “Train your own image classifier with Inception in Tensor- Flow,” https://research.googleblog.com/2016/03/train-your-ownimageclassifier-with.html, 2016.
[17] P. Sermanet, D. Eigen, X. Zhang, M. Mathieu, R. Fergus, and Y. Le- Cun, “Over feat: Integrated recognition, localization and detection using convolutional networks,” arXiv preprint arXiv:1312.6229, 2013.
[18] P. Wu, S. C. Hoi, H. Xia, P. Zhao, D. Wang, and C. Miao, “Online multimodal deep similarity learning with application to image retrieval,” in Proceedings of the 21st ACM international conference on Multimedia. ACM, 2013, pp. 153–162.
[19] S. Liu, Z. Song, G. Liu, C. Xu, H. Lu, and S. Yan, “Street-to shop: Cross-scenario clothing retrieval via parts alignment and auxiliary set,” in Computer Vision and Pattern Recognition (CVPR), 2012 IEEE Conference on. IEEE, 2012, pp. 3330–3337.
[20] K. Yamaguchi, M. H. Kiapour, L. E. Ortiz, and T. L. Berg, “Retrieving similar styles to parse clothing,” IEEE transactions on pattern analysis and machine intelligence, vol. 37, no. 5, pp. 1028–1040, 2015.
[21] J. Wan, P. Wu, S. C. Hoi, P. Zhao, X. Gao, D. Wang, Y. Zhang, and J. Li, “Online learning to rank for content-based image retrieval,” 2015.
[22] Pavani P., Prabha T.S. (2014) Content Based Image Retrieval Using Machine Learning Approach. In: Satapathy S., Udgata S., Biswal B. (eds) Proceedings of the International Conference on Frontiers of Intelligent Computing: Theory and Applications (FICTA) 2013. Advances in Intelligent Systems and Computing, vol 247. Springer, Cham.
[23] Saritha, R.R., Paul, V. & Kumar, P.G. Content based image retrieval using deep learning process. Cluster Comput.22, 4187–4200 (2019). https://doi.org/10.1007/s10586-018-1731-0